DET PSYKOLOGISKE FAKULTET

A Discussion of Theoretical Solutions to the Hard Problem of Consciousness

HOVEDOPPGAVE

profesjonsstudiet i psykologi

Martin Hvideberg Bratlid

Høst 2013

Veileder
Mark Price

Abstract

The problem of explaining subjectivity or sentience scientifically has been coined the 'Hard Problem of Consciousness'. Some postulate that an explanation is impossible with methods from contemporary cognitive science. In this paper, I select three of the most influential theories of consciousness based on a literature survey, and ask whether their respective theoretical solutions to the Hard Problem are supported by empirical evidence from research on visual consciousness, and more broadly whether empirical evidence favours the any of these approaches. I argue that the evidence does not warrant a conclusion, but that it favours viewing consciousness as an emergent property of global access/global neuronal networks rather than dependent upon specific brain areas.

Sammendrag

Å forklare subjektive opplevelser vitenskapelig refereres til som 'The Hard Problem of Consciousness'. Enkelte forskere og filosofer mener at en forklaring er umulig med metoder i fra dagens kognitive vitenskap. I denne oppgaven gjør jeg en litteraturundersøkelse og på bakgrunn av denne selekterer jeg tre av de mest inflytelsesrike teoriene om bevissthet. Videre undersøker jeg om noen av disse teoriene er støttet av forskning på visuell bevissthet. I oppgaven blir også bredere problemstillinger diskutert, som spørsmålet om hvorvidt forskningen ser ut til å favorisere en av tilnærmingene. Jeg argumenterer for at forskningen ikke støtter en konklusjon, men at det å se på bevissthet som en egenskap ved global tilgjengelighet eller globale nevronale nettverk er favorisert over det å tenke på bevissthet som avhengig av spesifikke hjerneområder

Table of Contents

The Hard Problem of Consciousness

Consciousness is a familiar concept to most of us, not so much in a scientific or philosophical sense, but in the day to day feeling of *being* conscious – i.e. having subjective experience. Block (1995) propose that one can distinguish between Access (A)-consciousness which is composed of whatever information is consciously accessible to a person at a given moment, and hence can be verbally reported to a researcher, and Phenomenal (P) -consciousness which is subjective and entails a 'what it is like to be' quality. Block (1995) claims that A-consciousness is what is studied by experimental psychology, while P-consciousness is largely ignored.

Chalmers (1995) distinguished between the easy problems of consciousness and the Hard Problem of consciousness. Easy problems are solvable using standard methods from cognitive science, either computational or neuronal. The Hard Problem on the other hand, is a problem that seems unsolvable when we apply methods from cognitive science. Chalmers mentions several examples of easy problems such as how we focus our attention, how we discriminate between stimuli or how we access internal states. Though the easy problems are far from solved, they are susceptible to explanations offered by cognitive models. The truly hard problem, according to Chalmers, is the problem of subjective experience. One of the most well-known ways of phrasing this particular problem is Nagel's (1974) 'what is it like to be a bat'. While it is possible to understand how the bat navigates by echolocation, it seems impossible to understand what it is like to orient oneself using this sensory modality. The point that both Nagel and Chalmers have in common, is that when brains process information there is also a subjective aspect to it, a 'what it is like to be' quality.

The mystery surrounding subjective experiences, also known as the Hard Problem of consciousness, primarily take two forms. The first one is the mystery of why subjective experience should exist at all. For example, the possibility of the existence of a human with all the capabilities that you would expect from a conscious human being, but operating as an automaton with no subjective experiences, could be argued (Chalmers, 2004). This philosophical-zombie thought experiment has been criticized (Carruthers, 2007; Dennett, 1988; 1991; 1995; Harnad, 1994), for being logically possible, but metaphysically impossible. Whether or not philosophical zombies are metaphysically possible, will in turn effect how the Hard Problem is resolved. If you accept the possibility of philosophical zombies, you must reject physicalism (Howell & Alter, 2009), which in turn will lead you to accept a form of dualism. I will discuss zombies and related thought experiments in relation to the theories of consciousness later on, as this represents an *a priori* position on the Hard Problem (Van Gulick, 2007).

The second part of the Hard Problem concerns how subjective experience is connected to human nervous systems. This is what Descartes (1641) was thinking about when he famously wrote that ‘res extensa‘ and ‘res cogitans‘ are two different substances that communicate through the pineal gland. There are now a number of studies that link consciousness to certain patterns of brain activity (Mormann & Koch, 2007; Rees et. al., 2002; Tononi & Koch, 2008). Finding the neural correlates of consciousness (NCC) is a promising project, but a number of questions still remain.

Thus the Hard Problem, from here referred to as HP- consists of two questions. A: How is it that at least some organisms (e.g. humans) have subjective experiences? B: What is the connection between subjectivity and the physical systems known as brains?

Assuming that having a subjective experience of an inner life and the outside world is an important feature of their lives for most people, it would be troubling if Block's (1995) claim that A-consciousness is the only thing that is studied in experimental psychology, turns out to be true. Where Block perceives a hole in the knowledge base created by experimental psychology, Chalmers (1995) goes further and asserts that explaining how we have subjective experiences is impossible with cognitive or biological models. It would therefore be of great value to find out whether or not Chalmers' proposistion is plausible. If this explanatory gap, which is addressed by the HP, is unbridgeable, we would need a paradigmatic shift to incorporate the inexplicable phenomenality of human experience.

Regardless of - or perhaps because of - the explanatory gap addressed by the HP, and the attention (Churchland & Churchland, 1998; Dennett, 1996b; Gray, 2004; Libet, 1996; Varela, 1996) and support (McGinn, 2000; Shear, 1999) that the HP has received, a number of psychological theories of consciousness have been formulated. However, the causal role of consciousness has been debated (e.g. Block, 1980; Chalmers, 2004). Views on this subject are influenced by whether one views subjective experience as something different from the physical world (Baars, 2007). Despite questions relating to the causal role of consciousness a multitude of philosophers and scientists have adopted a naturalistic approach (e.g. Crick & Koch, 1990, Dennett, 1991). In these approaches the empirical status of the concept of consciousness is considered acceptable, as long as it is defined carefully.

With a basis in this assumption, I discuss how some of the leading psychological theories of consciousness address the HP. Having introduced the HP and its central

relevance to a science of psychology in this section, the structure of the rest of the paper is as follows. The next section of this paper summarises a survey where I rank contemporary theories of consciousness in terms of how often they are mentioned in journal titles, keywords and abstracts. On this basis I select influential theories for further discussion in the paper. This is followed by a section where I discuss how some of the leading psychological theories of consciousness approach HP of consciousness, and a section where I discuss predictions that should follow from these approaches and ask whether these predictions can in principle be distinguished. In the second part of this paper, I ask whether work within one major area of consciousness research - namely work on visual consciousness - appears in practice to support one or other approach to the HP. The first section within this part presents a justification for focusing on visual consciousness as a test area for approaches to the HP, and provides an overview of work in this area. The next section then asks whether work in several areas of visual consciousness is consistent or inconsistent with proposed approaches to the HP, and whether it selectively supports one or other approach. Finally, I discuss the implications of the findings for the concept of the Hard Problem of consciousness.

Selection of Theories of Consciousness

Due to space considerations this paper will not address all the contemporary theories of consciousness which are relevant to the HP. Instead a subset of the most influential theories were selected on the basis of a literature search, as described below.

Selection Method for Theories

To narrow the initial search, a list of theories of consciousness was first based on the contemporary theories of consciousness discussed in the Blackwell Companion to Consciousness (Schneider & Velmans, 2007). As a second step this list was compared with the list of theories of consciousness in the Stanford Encyclopedia of Philosophy (Van Gulick, 2011). In this encyclopedia the theories are categorised as metaphysical theories if they are overarching theories that offer answers to the relationship between consciousness and the physical world. They are also categorised as specific theories if more detailed accounts of consciousness are given. This provided a framework for exclusion, as metaphysical theories were considered too broad for the purposes of this paper. Thus the following well known metaphysical theories, which offer explanations of the relationship between the physical world and consciousness rather than the specific mechanisms surrounding consciousness, were excluded: mysterianism (McGinn, 2004; Rowland, 2007), naturalistic dualism (Chalmers, 2007), biological naturalism (Searle, 2007) and reflexive monism (Velmans, 2007).

Representationalist theories (Seager & Bourget, 2007) are mainly concerned with the relationship between representations and phenomenality. Though not fitting the metaphysical theory category, representationalist theories refer to an overarching class of theories. Three higher-order theories (Carruthers, 2007), which are a specific class of representationalist theories, were included instead of the broader representationalist theories, due to the fact that they offer reductive explanations of phenomenal consciousness: the Higher-Order Thought Theory (Rosenthal, 1990), Inner-Sense theory (Armstrong, 1968; Lycan 1995), and Dual-Content theory (Carruthers, 2003)

Lastly, there was one set of specific theories, namely quantum mechanical theories (Hameroff & Penrose, 1996; Stapp, 2007) that was excluded. There are four reasons for this: Based on an interpretation of Gödel's incompleteness theorem, one central premise of the most well known quantum mechanical theory - the Penrose-Hameroff theory - is that mental processes cannot be algorithmical, and that consciousness arises due to quantum incoherence in the microtubules of the neurons (Penrose, 1994). Evidence from studies of anaesthesized subjects suggest that the proper functioning of microtubules are necessary for consciousness (Hameroff, 1998). On the other hand, there might be many things that are necessary for consciousness to arise. This and the critical reception of the argument based on Gödels theorem (Blackmore, 2010; Grush & Churchland, 1995; Putnam, 1994) suggest that quantum mechanical theories of consciousness are much debated, which makes it hard for those who are not acquainted with the conceptual language to make judgements about the validity and empirical status of the theories. Secondly, there are few predictions that have been made from proposing that consciousness arises from quantum incoherence in the microtubules, although the future possibility of comparing microtubule activity in conscious and unconscious states remains. Thirdly, quantum mechanical theories are vulnerable to a post hoc fallacy, because their central premise could be caricatured as saying that since consciousness seems so inexplicable, we must explain it using something inexplicable. In this way, quantum mechanical theories of consciousness are replacing the mystery of subjective experience with the mystery of quantum coherence. Lastly, the quantum theory approach to consciousness falls beyond the conventional cognitive neuroscience focus of the current paper.

The theories that were included were the Multiple Drafts Model (Dennett, 1991) Global Workspace Theory (Baars, 1988), Higher-Order Thought Theory (Rosenthal, 1990), Inner-Sense Theory (Armstrong, 1968; Lycan, 1995), Dual-Content Theory (Carruthers, 2003) Intermediate Level Theory (Prinz, 2007), and Information Integration Theory (Tononi, 2007).

Some of the theories mentioned above also have other names that are used interchangeably those used in this paper. For example, Dual-Content Theory is sometimes referred to in the literature as Dispositionalist Higher-Order Thought Theory and Intermediate Level Theory is sometimes referred to as Attended-Intermediate-Representation (AIR)-theory. When searching for theories with multiple names, all the known labels for that theory were used.

Selection Method for Journals

A preliminary search was conducted within the 50 top ranking psychology journals in terms of JCR impact rating (University of Massachusetts Amherst Libraries, 2013). In addition to these, two more journals were added based on the mind-science foundations list of journals frequently reporting scientific findings in consciousness research (Mind-Science Foundation, 2013). These were considered high impact journals within the consciousness research fields. The initial list of journals was thus based mostly on JCR impact rating, but with two journals that were included due to the endorsement from mind-science foundation. The names of the selected theories were

used as search terms that could appear anywhere in an article, and there was no publication date constraint.

The preliminary survey revealed that Behavioral and Brain Sciences had the most hits (120), followed by Journal of Consciousness Studies (99), Consciousness and Cognition (96), Trends in Cognitive sciences (20), Neuropsychologia (11), Cognition (8), Cognitive Psychology (7), Journal of Cognitive Neuroscience (3), Annual Review of Psychology (2), Psychological Review (2), Journal of Psychiatric Research (2), Psychological Bulletin (1), Personality and Social Psychology Review (1), Journal of Child Psychology and Psychiatry (1), Journal of Experimental Psychology - General (1), Psychological Medicine (1), Current opinion in Psychiatry (1), Developmental psychology (1). The remaining 34 journals had no hits for any of the theories.

The seven journals with the most hits in the preliminary search were selected for the final survey: i.e., Consciousness and Cognition, Trends in Cognitive Sciences, Journal of Consciousness Studies, Behavioral and Brain Science, Neuropsychologia, Cognition and Cognitive Psychology. The final survey was conducted by searching for articles that mentioned the target theories in their abstract, keywords or titles.

Results

The Global Workspace Theory was mentioned in most abstracts, titles and/or keywords (28), followed by the Higher-Order Thought Theory (10), the Multiple Drafts Model (6), the information integration theory (4) and the Intermediate Level theory (1).

The Inner-Sense Theory and Dual-Content Theory were not mentioned in any abstracts, titles or keywords in these journals. The journal of consciousness studies has the most hits (19), followed by Consciousness and Cognition (14) Behavioral and Brain sciences (9), Trends in Cognitive sciences (5), Cognition (2) and Cognitive Psychology (1). Neuropsychologia had no hits for any of the theories (see Table 1)

Journal	MDM	GWT	HOT	ILT	IIT	IST	DCT	Sum
Consciousness and Cognition	2	7	1	1	3	0	0	14
Trends in cognitive sciences	0	4	1	0	0	0	0	5
Journal of consciousness studies	2	12	4	0	1	0	0	19
Neuropsychologia	0	0	0	0	0	0	0	0
Cognition	0	2	0	0	0	0	0	2
Behavioral and Brain Sciences	2	4	3	0	0	0	0	9
Cognitive Psychology	0	0	1	0	0	0	0	1
Sum	6	28	10	1	4	0	0	49

Table 1. Number of times that the theory appeared in an abstract, title and/or keyword. From the left MDM: Multiple Drafts Model, GWT: Global workspace theory, HOT: Higher-Order Thought Theory, ILT: Intermediate Level Theory, IIT: Information Integration Theory, IST: Inner-Sense Theory, DCT: Dual-Content Theory

Implications

The Global Workspace Theory appeared three times as often as the runner up, the Higher-Order Thought theory in abstracts, keywords and titles, with the Multiple Drafts Model a further five hits behind Higher-Order Thought Theory. The hits from the Behavioral and Brain Sciences, which is the journal in the survey with the highest JCR-impact factor, show the same tendency as the survey as a whole. Including the results from the other journals with a high JCR-impact factor, Trends in Cognitive science, Cognition and Cognitive Psychology, the ranking of the theories held constant. These results suggest that the hits from Journal of Consciousness studies and Consciousness and Cognition, which are significant journals in their field, correspond with hits from the higher ranking journals, indicating that the findings from this survey reflect the mainstream tendency.

Based on these results, the Global Workspace Theory (GWT), the Higher-Order Thought Theory (HOT), and the Multiple Drafts Model (MDM), were selected as the leading theories for the focus of this paper, both in offering a specific perspective on the HP, and in being evaluated in light of empirical evidence from research on visual consciousness. I will briefly outline these theories in the following section, and summarise as well as compare their empirical predictions in the next.

The Global Workspace Theory

The GWT describes conscious events as happening in a metaphorical ‘theatre of consciousness’ (Blackmore, 2010). This view emphasises the large difference between the very few items that are in consciousness, and the vast amount of unconscious neural processes that are going on at any time (Baars, 2007). The perspective from GWT holds that this can be thought of in terms of a theatre where consciousness acts as the spotlight on the stage, which is directed to different places by attentional processes. The conscious spotlight is surrounded by events at the fringes that are only vaguely conscious. The audience sits in the unconscious darkness and receives information from whatever is lit up by the unconscious spotlight, and behind the stage there are various unconscious processes that shape the events happening in the spotlight. This ‘theatre metaphor‘ is based on a blackboard architecture, where consciousness is seen as a limited capacity workspace with flexibility to solve novel problems, and the unconscious processes are thought of as more specialized (Baars, 1988). The use of the metaphor ‘Theatre of Consciousness’, ‘Global Workspace’, or ‘blackboard architecture’ in describing consciousness, does not imply that consciousness must be localized to one area of the brain where everything comes together (Baars, 2002).

The GWT provides a functional explanation for conscious experience, in viewing consciousness as a limited-capacity global workspace (Baars, 1988; 2002; 2007). The main assumptions are that while unconscious mental processes mainly are specialized, conscious processes are not. Consciousness is viewed as the blackboard of the mind, where bits of information can be held, manipulated, and fed back to more specialized unconscious information processing modules. This provides a functional answer to question A in the HP (How is it that some organisms have subjective

experiences?). It does not give us an answer to question B (What is the connection between subjectivity and the physical systems known as brains?).

The Neuronal Global Workspace Theory (Dehaene & Naccache, 2001) however, is similar to the GWT in the sense that it has a basis in the global workspace architecture from the GWT. In line with this it is considered a submodel of this theory (Baars, 2007). It also assumes the existence of unconscious processors that compete for access to the limited capacity of the global workspace, but the neuronal GWT also focuses on the neural foundation of consciousness. According to this view, subjective experience is dependent upon long range circuits involving the prefrontal cortex, anterior cingulate and connecting areas (Dehaene et. al., 2006). If information enters into connection with the long range circuitry it can be widely broadcasted to other brain areas, thus creating a global availability that results in the possibility of verbal or non-verbal report and is experienced as a conscious state. The neuronal GWT hypothesize that subjective experience results from the creation of a global availability of information due to long range circuitry in the brain. Thus the GWT provides a possible answer to question A and the neuronal GWT a possible answer to question B.

GWT is an umbrella term for many Global Workspace sub-models. Some of these are outlined by Baars (1988; 2007), and some by other authors (e.g. Dehaene & Changeux, 2000; Dehaene, et. al., 2001; Dehaene, et. al. 2003; Shanahan and Baars, 2005). While the sub-models make specific predictions, there are some general theoretical predictions that are implied by a global workspace architecture. The most important one is the conscious access hypothesis, where it is theorized that consciousness is a function that enables the integration and mobilization of separate

brain functions that are otherwise independent (Baars, 2002). The predictions that follow from the GWT will outlined in the next part of this paper, after the other theories have been descrbed. Thus comparing the different predictions as well as describing them.

The Higher Order Thought Theory

There are a variety of HOT theories (Carruthers, 2007), but only the actualist-HOT theory, defended by Rosenthal (1990), had enough hits in the survey to be included in this paper. Therefore, when I refer to HOT in this paper, I will be referring to Rosenthal's version. According this version of HOT, a person is conscious of her mental states if she has a higher-order thought about that mental state (Rosenthal, 1990; Van Gullick, 2011).

Before I outline the predictions that can be drawn from HOT, I will briefly describe the assumptions that form the basis of the theory, starting with the distinctions that are drawn in HOT between different types of consciousness. *Creature consciousness* is the first type, and it is posessed by creatures who are awake and sentient. Assuming that wakefulness and sentience are biological features of an organism, this does not seem to be something that directly addresses the HP (Rosenthal, 1990). The second type of consciousness is *transitive consciousness*, which is consciousness that takes an object, e.g. consciousness *of* something. Thirdly, *state consciousness* or *intransitive consciousness* is a type of consciousness that involves a meta-mentality. In this sense there is a difference between wanting, for example, an apple and being conscious of one's wanting an apple (Droege, 2005). It is this type of

consciousness that is primarily addressed in HOT theory and from now on and as long as I am discussing HOT, when I refer to consciousness I imply state consciousness.

As they are based on the distinction between transitive and state consciousness, all HOTs - including Rosenthal's, are based on the transivity principle (TP) (Rosenthal, 2008; Rosenthal & Weisberg, 2008), a principle stating that a conscious state is experienced as conscious when you are aware of the state in a suitable way. At first glance, this might sound tautological and simplistic, but it does have some important implications. It implies that if a person has a thought, perception or a feeling but is wholly unaware of it, it would not be a conscious state. Thus, a conscious state according to the TP would be dependent on the persons access to that state.

As mentioned earlier, HOT assumes that a higher-order thought about a mental state is sufficient to make that mental state conscious. Still, there are some important factors that must be present for a person to be conscious of her mental states. As well as stating these factors, I will give an example for each of them as this will make them more tangible, hopefully easing the transition into the section where the predictions are discussed. First of all, state consciousness is dependent upon a sense of immediacy (Byrne, 1997). When a person is conscious of her wanting an apple, she is not inferring it from other information nor is she observing it. She is not thinking 'I feel hungry, and I have not eaten anything with a lot of dietary fibres in a while, I should grab an apple'. There are no conscious observations or presumptions that mediate between a mental state and the higher-order thought. Secondly, a mental state is conscious by having a representational relation to a higher-order thought (Droege, 2005). It is not conscious due to an intrinsic property. When the person from the previous example is conscious of

her hunger for apples, it is conscious because of the relation the representation of her hunger for apples has to a higher-order thought. Thirdly, a higher-order thought must occur at approximately the same time as the mental state it represents, and it must assert, in a strong manner, that you are in a particular mental state. For the woman with the hunger for apples to be conscious of this aspect of her state of being, the higher-order thought about her hunger for apples must occur at roughly the same time as her bodily hunger occurs. And it cannot be her speculating or hoping that what she is experiencing is a hunger for apples it has to be an assertion of her hunger for apples. Fourth, it should be noted that a higher order thought is not equivalent to an intentional introspective attempt on the part of the subject. A higher-order thought about the higher-order thought would be needed if a higher-order thought was to be available for introspection (Rosenthal, 2000). Lastly, a minimal self concept is required for a higher-order thought to occur, due to the self-reference that is made in higher-order thought. The woman with the craving for apples would have to know that she was the one with this want, and not anyone else.

HOT is different from most other theories of consciousness due to its simplicity. It does not postulate any special organs for consciousness, nor does it postulate any particular function for consciousness, like adopting a stance where conscious processing has an advantage over unconscious processing, for example when handling novel information. However, Rosenthal (2008, pp. 830-831) explicitly writes that this conclusion about the function of consciousness does not imply epiphenomenalism. He goes on to write that the absence of function does not mean that consciousness is devoid of a causal role on other psychological processes, but that this causal role is too limited,

diverse or neutral with regards to the interests of the organism to have any significant effect on the function of consciousness. In line with this, HOT appears to be agnostic with regard to the function of consciousness, as Lau and Rosenthal (2011) writes:

«The higher-order view, by contrast, is neutral about whether conscious awareness adds significant utility or immediate impact on behavior and task performance. This is because the view assumes that task performance in most perceptual and cognitive tasks depends mainly on first-order rather than higher-order representations. Because conscious awareness can differ even if all first-order representations remain completely unchanged, such awareness itself might serve little function» (Lau & Rosenthal, 2011, pp. 366-367)

Avoiding any predictions about the function of consciousness makes the theory more defensible in Rosenthal's view (Lau & Rosenthal 2011). Coherent with this line of reasoning, HOT simply purports to outline the pre-requisites and mechanisms that makes something conscious. If we were to return to question A of the HP (How is it that at least some organisms, e.g. humans, have subjective experiences?), the answer from HOT would simply be, because humans have higher-order thoughts with a representational relation to their experiences. If we return to question B of the HP (What is the connection between subjectivity and brains?), HOT predicts that changed awareness is associated with changes in the activity of the prefrontal and parietal regions of the brain (Lau & Rosenthal, 2011), though it is stressed that not any changes in activity would be sufficient for change in awareness and that higher-order

representation probably represents a small sub-set of such an activity. HOT's stance on the HP is thus more descriptive than the GWT's.

The Multiple Drafts Model

Dennett (1991) presents two views of consciousness, one of which disputes the approaches that many scientists and philosophers have adopted in trying to account for consciousness. The other view is his own perspective on consciousness, the Multiple Drafts Model (MDM), which draws on models from evolutionary theory, as well as psychology, anthropology, biology, neuroscience and artificial intelligence research. The proposal of the MDM is that our brains consist of a number of multiple channels with specialist circuits that try, in disorganized parallelity (Dennett, 1991, pp. 253-254), to do their different tasks, creating "multiple drafts" as they go. These drafts are fragments of narrative, and while most of the drafts are short lived, and only modulate current activity, some are promoted to other useful roles, by a virtual machine in the brain, which operates as a serial processor. This virtual machine arises from a coalition of specialist circuits, and its seriality is not hard wired in the brain (Dennett, 1991, pp. 253-254), but is an effect of the succession of coalitions of specialist circuits coming together. The specialised circuits outside of this virtual machine normally operate as a parallel processor, as parallel processes are seen as an important attribute of the brain.

Many of the specialist circuits are shared with our pre-historic ancestors, but are opportunistically enlisted to new roles, even if they originally evolved because they increased the likelihood of detecting environmental hazards, like spoiled food or predators. The reason why the result is not more anarchic, is that the central tendency

that is imposed on this activity is a product of evolution, culture or individual microhabits and hence a design feature (Dennett, 1991, pp. 253-254). Hence, some of the design features are innate, and thus shared with other animals, other parts of the design features are augmented, undermined or mediated by microhabits developed by the individual and pre-designed microhabits that are culturally transmitted. In later work the metaphor of multiple drafts has been replaced with the metaphor of "fame in the brain", but the underlying point remains the same. Consciousness is not a datable event, like something being transmitted via a medium (e.g. televised), it is a gradual process, presicely like fame (Dennett, 1996a).

Central to the MDM (Dennett, 1991) is also that we need to think differently about consciousness, and to be able to do this, we need to get some of the common misunderstandings about consciousness out of the way: One of these is the powerful day to day intuition about what Dennett calls a Cartesian Theatre. This is the familiar idea that various functional modules in our brains like perception, working memory and attention perform the basic legwork of making information understandable through processing before they *present the information to our conscious selves*. Dennett (1991) lists several reasons why this cannot be true, one of the most persuasive ones being that there are several problems with trying to locate the point of view of the individual, within the individual.

Other powerful intuitions about consciousness include the intuition of a singular narrative stream, and the intuition of the existence of an author (Dennett, 1991). The intuition about the existence of a singular narrative stream arises when the parallell processes are probed in one way or the other, for example, the contents of consciousness

can be probed as we engage in conversations about our experiences or when we make judgements about what our experiences are (Blackmore, 2010). From these everyday activities the illusion that there is a single stream of consciousness arises. The intuition about the existence of an author of this single stream of consciousness is according to Dennett, also an illusion. This illusion comes into play when the contents of consciousness are made clearer by probing the parallel processes at various points. This too happens as we engage in our everyday routine of talking about our experiences and making up our minds about what they are and how we should act on them.

Both the intuition about a singular stream and an author of consciousness are illusions, but the argument so far does make references to 'probing parallel processes'. Due to this an observer can be inferred, but Dennett escapes the problem of defining a point of view within the observer, by claiming that the observer is a skein of narratives (Blackmore, 2010) - or a «centre of narrative gravity» in his own words, thus drawing on the concepts he developed in *«The Intentional Stance»* (Dennett, 1989).

The MDM is a functionalist model. Dennett (1991) claims that any model that does not solve the problems of what functions consciousness performs, is not a satisfying model. To question A (How is it that at least some organisms (e.g. humans) have subjective experiences?) of the HP, it is implied that the subjective qualities of experience only accessible to the individual must be analyzed in a functional manner and broken down in complex dispositional traits distributed in space and time in the brain (Dennett, 2001). In other words, as long as you think about consciousness in the right way, you will come to see that there is no HP, which is in line with previous critique of the concept phenomenal consciousness or qualia (Dennett, 1988). Turning to

question B of the HP (What is the connection between subjectivity and the physical systems known as brains?), the view from MDM is that consciousness is like fame in the brain (Dennett, 1996a). At any given time the information that is processed by the largest population of specialist circuits is conscious. Viewing the answers to question A and B, in relation to each other, the MDM holds that the HP is an exaggerated way of phrasing the problem of subjectivity, and that consciousness is a 'virtual machine' (Dennett, 1991, pp. 253) that shapes information processing in the mind.

The MDM was intended to be a novel way of understanding consciousness, though Dennett (1991, pp 254-255) concedes that it borrows heavily from other theories of consciousness. The idea that there is no homunculus in the brain was endorsed by many scientists and philosophers at the time, and so was the idea that the brain works as a parallel processor (Schneider, 2007). Though this might be the case, the MDM was simply intended to be a sketch of how a model of consciousness could be conceived without a Cartesian Theatre (Dennett, 1991). However, the model has a set of specific predictions built into it, but any specific model that honors the key propositions of the MDM could in principle be called a Multiple Drafts Model (Dennett & Akin, 2008)

Predictions from the Theories of Consciousness

As shown in table 2, there are some similarities between the HOT, GWT and MDM. All of the theories claim that conscious awareness in some ways involves prefrontal regions (Baars, 2002; Dennett, 1991; Lau & Rosenthal, 2011). Though the details surrounding this are different between the theories. The HOT predicts that

conscious awareness is dependent upot prefrontal regions, as this is where the higher-order thoughts arise. Both the GWT and MDM also predict the involvement of prefrontal regions in consciousnes, but not a dependence upon these regions. Hence they are differentiated between, as the prediction that is made by the GWT and the MDM, is that consciousness is an emergent property of global neuronal networks (Baars, 2002; Dennett, 2001).

Another difference - which also is between the GWT and MDM on the one hand and the HOT on the other, concerns whether they assign a functional role to consciousness. The HOT does not assign any functional role to consciousness (Lau & Rosenthal, 2011; Rosenthal, 2008). The GWT and MDM however, attribute a functional role to consciousness. This functional role is the integration and mobilisation of various brain resources. In the MDM this is phrased as a ‘coalition of specialist circuits coming together’ (Dennett, 1991, pp. 253) whereas in the GWT this is phrased as ‘a global workspace’ or ‘blackboard architecture’ (Baars, 1988; 2002; 2007).

Similar for all of the theories is that they propose the self as the dominant context of experience. In the HOT, this self reference is theorized to be related to an innate capacity, a «theory of mind» turned inwards (Droege, 2005; Rosenthal, 1990). Comparable to this, the GWT holds that the self is a group of ‘executive interpreters’ located in the frontal cortex (Baars, 1988;2007). While the proposition from the MDM is that the self is an abstraction, created by a vast number of attributions and interpretations that have constructed the biography of a person. This abstraction is referred to as ‘the centre of narrative gravity (Dennett, 1991, pp. 427-428) and forms the central context around new experiences that are attributed and interpreted.

Though the theories share the prediction about self reference, both the MDM and the GWT differ from the HOT in predicting that consciousness is a graded phenomenon (Baars, 1988; 2007; Dennett, 1991; Dennett & Akins, 2008). With basis in the MDM and GWT it would make no sense to ask *when* something became conscious. The catchphrase that consciousness is similar to 'fame in the brain' (Dennett, 1996) alludes to this point. This catchphrase also points to the sixth theoretical prediction, which the MDM and GWT also share. That consciousness is a spatially and temporally distributed set of processes, which again hints at consciousness arising due to a mobilization and integration of brain resources. HOT however, with its emphasis on prefrontal areas (Lau & Rosenthal, 2011) and the mechanism of something becoming conscious when a higher-order thought has a representational relation to a mental state (Rosenthal, 1990), does not make predictions regarding a spatial and temporal distribution of consciousness. Neither does it predict that consciousness is a graded phenomenon; a higher-order thought with a representational relation to a mental state either exists, or it does not.

Lastly, the HOT makes one more theoretical prediction that it does not share with the MDM and GWT: That consciousness emerges as a result of a representational relation between a higher-order thought, and a mental state (Rosenthal, 1990). Higher-order thoughts are located in the prefrontal area of the brain. (Lau & Rosenthal, 2011)

	GWT	**HOT**	**MDM**
Consciousness is dependent upon prefrontal areas	-	x	-
Conscious awareness depends on a large global network of neuronal clusters in several areas of the brain.	x	-	x
Functional role: Global access in contrast to specialised unconscious processes consciousness	x	-	x
Self-reference: The self as the dominant context of experience	x	x	x
Consciousness is a graded phenomenon	x	-	x
Consciousness is a spatially and temporally distributed set of mental processes	x	-	x
Consciousness arises from a representational relation.	-	x	-

Table 2. Comparison of Theoretical predictions from GWT, HOT and MDM. X indicates positive, - indicates negative. Based on the following sources: The predictions drawn from the MDM are based on Dennett (1991, pp. 253-254, 464-468; Dennett, 1996) and Dennett & Akins (2008). Whereas the predictions made by the HOT are on based articles (Byrne 1997; Lau & Rosenthal 2011; Rosenthal 2000) and relevant book chapters (Blackmore 2010; Carruthers, 2007; Droege, 2005; Rosenthal, 1990; Van Gulick, 2011). Lastly, the predictions that the GWT make are from articles (Baars, 2002; Dehaene and Changeux, 2000; Dehaene et. al., 2001; Dehaene & Naccache, 2001; Dehaene et. al., 2003; Dehaene et. al., 2006) Franklin, 2003; Shanahan and Baars, 2005), books (Baars, 1988) and relevant book chapters (Baars, 2007)

The theoretical predictions in table 2. have two major themes which directly feeds back to the HP, one concerns question A (e.g. how is it that some organisms have subjective experiences?) of the HP. This is touched upon by the question of whether consciousness is assigned a functional role. On this question the answer from the MDM and the GWT is 'yes', whereas HOT does not propose any function for consciousness (Rosenthal, 2008). However, if consciousness should be shown to have a function, this would not in itself falsify the HOT as HOT is agnostic towards the function of consciousness (Lau & Rosenthal, 2011).

The other theme in table 2. is related to question B of the HP (What is the connection between subjectivity and the physical systems known as brains?), and concerns how consciousness is related to activity in the brain. In contrast to the former theme where the GWT and MDM made a stronger claim. The stronger claim with regards to consciousness and brain activity is made by the HOT, which proposes that activity in the prefrontal areas are a prerequisite for consciousness (Lau & Rosenthal, 2011). The MDM and GWT however, suggest that consciousness arises due to global neuronal networks (Baars, 2007; Dennett, 1991) that likely also involves frontal areas. The GWT for example does hold that the 'self' is the dominant context of experience, and that the 'self' is a cluster of 'executive interpreters' (Baars, 1988) located in the frontal cortex. Something that implies that a degree of frontal activity is predicted in conscious experience. However, this is essentially different from proposing that frontal activity is a prerequisite for conscious experience.

Summarised there are four competing hypothesis directly linked to the two aspects of the HP. For simplicity I shall use the letter A to refer to the predictions

relavant to question A of the HP, and the letter B to refer to predictions relevant to question B of the HP:

A1) Consciousness has a function: Mobilisation and integration of brain resources (MDM/GWT).

A2) Consciousness probably does not have a function (HOT).

B1) Consciousness is dependent upon activity in the prefrontal cortex (HOT).

B2) Consciousness is an emergent property of global neuronal networks (GWT/MDM).

These four competing hypothesis which directly adress the HP are the focus of the rest of this paper. In the next section I will provide a justification for chosing visual consciousness as a test field, before I summarise relevant empirical evidence and discuss it in relation to the competing hypothesis.

A Justification for Using Visual Consciousness as a Test Field

Visual consciousness is by many scientists viewed as a field that can inform the understanding of consciousness (e.g. Engel, et. al., 1999; Sheinberg & Logothetis 1997; Zeki & Bartel, 1999). The advantages that this area has over others include: (1) Awareness being linked to an outside stimulus with a specific set of reactions associated

with it over a series of trials; (2) Visual experience entails a 'what it is like' quality - indeed, it has been used by philosophers in thought experiments (e.g. inverted qualia, and absent qualia thought experiments, Block, 1980); (3) Visual systems among many species of vertebrates are relatively similar which makes animal models possible (As opposed to studying intentionality or abstract thinking).; (4) Vision is assumed to be central to how we orient ourselves and form spatial maps over the environment, so it arguably plays a larger role here than our auditory, olfactory and tactile sensory systems; (5) Lastly, with the current methodology, the visual consciousness field is one of the few fields where consciousness explicitly is treated as a variable.

This last point is particularly important when comparing the competing hypothesis. The evidence reviewed here will have to shed light on the relationship between function and consciousness to inform question A of the HP. It will also have to illuminate the relationship between consciousness and neural activity to have any bearing on question B of the HP.

Structure for Discussion of Findings

Before entering a more detailed discussion of relevant findings from studies of visual consciousness, I clarify one basic structure of my discussion.

First, to address the relationship between function and consciousness I draw on evidence from studies of neurological patients with lesions to their striate cortices that have blind fields corresponding to the field that is retinotopically mapped by the

lesioned area (Cowey & Stoerig, 1991). The phenomenon of blindsight refers to these patients being able, for example, to grasp and manipulate objects in their subjectively blind field (Weiskrantz & Weiskrantz, 1986). This phenomenon will be discussed in relation to the competing hypothesis addressing question A of the HP.

Second, in relation to the competing hypothesis addressing question B of the HP, I address studies comparing effects of consciously and unconsciously processed stimuli. In the method of 'contrastive analysis' (Baars, 2007), experimenters compare the difference between the effects of a stimuli that is consciously processed and the same stimuli when it is unconsciously processed. Studies comparing effects of consciously and unconsciously processed stimuli, can be found for example in visual masking (Price, 2001) and the binocular rivalry paradigm (Blake & Tong, 2008).

Consciousness and Function

In this section I will review the evidence from the area of visual consciousness which relates to the competing hypothesis about the functions of consciousness, i.e. A1. Consciousness has the function of mobilisation and integration of brain resources - MDM/GWT; A2. Consciousness probably does not have a function - HOT. However, I will first outline some of the claims that have been made about consciousness and function, as this will form the basis of the discussion of the evidence in this section. Some of these claims are based on thought experiments and some are interpretations of evidence. First, some thought experiments.

The inverted qualia thought experiment, sometimes referred to as the inverted spectrum thought experiment (Block, 1980), is a refined version of the profound question that even children sometimes ask, ‘is my red the same as yours?’. Qualia is a term philosophers often use to describe phenomenal experiences, and inverted qualia refers to an imagined scenario where the qualitative aspects of for example colour are inverted. So that - in this scenario - when light with a low wavelength frequency hits your retina, you will not experience red, but blue - which normally is associated with high wavelength frequencies. However, there would be no possible way of you knowing that you experienced blue. This would be your red, and it would make no functional difference. If inverted qualia are possible, there would be problems associated with ascribing functional properties to mental states.

The absent qualia thought experiment (Block, 1980) is similar, but the argument proceeds differently. It assumes the possibility of a state, for example my experience of hunger, to be functionally the same as another person’s experience of hunger. However, the other person’s experience completely lacks the qualitative aspects of my experience of hunger. This imagined person who lacks the qualitative aspects of experience, is similar to what is referred to as a ‘philosophical zombie’ (Chalmers, 2004). A person who can engage in all the complex behaviors you would expect from a conscious human, while lacking any conscious experience.

There are multitudes of thought experiments like these, that are continuously discussed, refuted and elaborated on in philosophical journals. However, the metaphysical possibility of the thought experiments mentioned here are disputed (Carruthers, 2007; Dennett, 1988; 1991; 1995; Harnad, 1994), as thought experiments

like these often represent a priori views on empirical questions (Van Gulick, 2007). Neurological conditions that scientists and philosophers link to the themes from these thought experiments are therefore of great interest.

A condition that has been linked to the absent qualia and philosophical zombie thought experiment is blindsight. This phenomenon refers to the fact that many patients with lesions to their striate cortices respond to visual stimuli presented in the part of the visual field that is retinotopically mapped by the lesioned cortex even though they do not consciously see those stimuli (Weiskrantz & Weiskrantz, 1986). Lau and Passingham (2006) wrote that this phenomenon is a dissociation between function and visual consciousness. McGinn (1991) is in line with this and interpreted the condition as implying that a person can function visually as well as any other, without being conscious of it. Persaud and Lau (2008) go further and present findings from an N=1 study, which they claim are evidence for absent qualia in the blind field of a patient with a lesion to the occipital lobe.

This last point, about absent qualia, is relevant only as long as the evidence points to the absence of any functional differences between blindsight and perception as it occurs in healthy brains. At first glance, the reports from the neurological studies are impressive. Ramachandran and Blakeslee (1998) for example, reported that patients were unable to tell how many fingers the examining doctor held up as long as it fell into the field that was retinotopically mapped by the lesioned area. Despite this the patients could, with above chance accuracy, pick up objects placed in the area that is retinotopically mapped by the lesioned area, and move them in ways that imply that they are to some extent aware of their visual environment. Similarly, the same authors

describe patients with larger lesions to the striate cortex, rendering them completely cortically blind. This group can also pick up and manipulate objects, which is surprising considering the fact that these patients report that they have no visual experience of the outside world.

Before discussing whether this condition is something resembling a real world philosophical zombie, I will outline the mechanisms that have been suggested to account for the phenomenon. A number of scientists have proposed an explanation in terms of a phylogenetically old neural pathway, and a newer dorsal and ventral neural pathway (Milner & Goodale, 1995; Mishkin & Ungerleider, 1983; Ramachandran & Blakeslee, 1998). The phylogenetically old pathway goes from the eye to superior colliculus in the brain stem, and from there to the cortex, most notably the parietal areas. This pathway is associated with orienting reactions (Dean, et. al., 1989; Sahibzada, et. al., 1986). The ventral and dorsal pathways are newer, in terms of an evolutionary time scale, and project from the eye to the lateral geniculate nucleus, in the thalamus - which serves as a relay station - and from there to the primary visual cortex where they branch into a ventral and dorsal stream. This dissociation is supported by multiple chains of evidence, both from studies of neurological patients (Hodges, et. al., 1998) and from fMRI studies (Culham, et. al., 2003; Valyear, et. al., 2006; Shmuelof & Zohary, 2005). The dorsal stream continues from the visual cortex towards the parietal cortex and is mainly concerned with grasping, navigating and other spatial functions. It has therefore been referred to as the ‘how’ pathway (Ramachandran & Blakeslee, 1998, pp. 56-57). The ventral stream proceeds from the visual cortex to the temporal lobes, and is primarily involved in object recognition. Due to this association, the ventral

stream has been referred to as the ‘what‘ pathway (Ramachandran & Blakeslee, 1998, pp. 56-57).

Assuming that the normal input to the ‘how’ and ‘what‘ pathways is severely impaired in cortically blind patients, due to the damage to the striate cortex (Weiskrantz & Weiskrantz, 1986), scientists have suggested that blindsight occurs partly because the phylogenetically old subcortical pathway remains intact (Lyon et. al., 2010; Morris, et. al., 1999), hence making orienting responses possible. Others have proposed that blindsight could be due to residual function within the primary visual pathway (Fendrich, et. al., 1992). One variety of this is the suggestion that some patients with this brain lesion might have parts of their ‘how’ pathways intact (Milner, 1998). These scientists point to the existence of another neurological condition called ‘Balint’s syndrome’ which is caused by a bilateral lesion to the parietal cortex, hence also a damage to the ‘how’ pathway (Ramachandran & Blakeslee, 1998). These patients can identify objects, but experience a form of tunnel vision limited to the area lined up with their fovea and often miss by several decimeters when they are asked to focus on a visual stimuli, using long periods of time to do something that normal subjects do effortlessly. This severe impairment is presented as the opposite of blindsight as the navigation aspects of vision are impaired whereas the knowledge and object recognition aspects remains.

Applying the neurological evidence reviewed here to the dissociation between visual consciousness and performance (Lau & Passingham, 2006; McGinn, 1991; Persaud & Lau, 2008), it would appear as if the authors proposing this dissociation are of the opinion that a lesion to the striate cortex will remove subjective experience from

the area that is retinotopically mapped by the lesion. However, another line of evidence from research on the common blind spot, and scotomas arising due to lesions in the striate cortex, points to the scotomas being perceived as filled in with color, texture and form (Ramachandran et. al., 1993; Ramachandran & Blakeslee, 1998). It is hard to imagine a scenario where these perceptions do not have an experiential dimension. Not only do the perceptions have phenomenality, they also have functional qualities, as the 'filling in' occurs in accordance with the background (Ramachandran & Gregory, 1991). Experiments on macaque monkeys show that this effect also can be seen on a neuronal level. De Weerd et. al. (1995) reported increased activity in extrastriate neurons with receptive fields covering the blindspot, corresponding to the effect of the gradual filling in. While another study has reported that some neurons in the striate cortex respond selectively to large stimuli covering the blind spot (Komatsu, et. al., 2000). Thus indicating that some neurons are essential for transmitting information concerning large uniform structures covering the blind spot, as well as the absence of smaller stimuli in this part of the receptive field.

In summary, it appears that the effect of filling in the blind field has both a phenomenal and functional aspect, which can be seen at both a behavioral and a neuronal level. This diverges from the finding reported by Persaud and Lau (2008), about the absence of qualia in the blind field. Considering the methods that they used, this is not surprising. They showed their participant common definitions of qualia, and asked him whether he felt he experienced those in his blind field. According to the Stanford Encyclopedia of Philosophy, there are at least four different definitions of qualia (Tye, 2013). Some of them are broad and easy to relate to, others are not.

Assuming that the participant is a non-philosopher, this is a possible confounding factor. In line with this Persaud and Lau (2008) reported that the participant failed to understand the difference between 'ineffable' and 'private', a disadvantage he shares with the author of the current paper. When the methodology is taken into consideration, and the finding is viewed in context with other evidence, it appears as if this is a result of a fuzzy concept being used as if it was a well defined one.

Having established reservations towards the evidence for absent qualia in the blindfield, the next question is whether blindsight provides evidence for a dissociation between visual consciousness and function, as was proposed by McGinn (1991) and Lau and Passingham (2006). The evidence presented in neurological studies sound convincing, but there is one important caveat. For the patients with lesions in their striate cortices to be able to manipulate objects in their blind field, they need verbal prompting in so-called forced choice behavioral tests (Cowey & Stoerig, 1991). They are reluctant to make judgements about visual stimuli in their blindfields, and they perform poorer than subjects with healthy brains in these tasks (Ramachandran & Blakeslee, 1998). Thus, the dissociation between visual consciousness and performance can be more precisely thought of as a dissociation between the experience of vision and a rudimentary ability to act adaptively. The philosophical zombie of real life presents more ambiguity than the one imagined by philosophers, and is not as applicable to thought experiments. Which in turn means that this phenomenon does not provide any evidence for thinking in terms of a functional and phenomenal dissociation. On the contrary, it provides evidence that losing the experience of visual consciousness has real implications for the functional aspects of vision.

Consciousness and Neural Activity

In this section I shall review evidence bearing an impact on question B of the HP, i.e., what is the connection between subjectivity and the physical systems known as brains? The competing hypothesis derived from the theories are, B1: Consciousness is dependent upon activity in the prefrontal cortex (HOT), and B2: Consciousness is an emergent property of global neuronal networks (GWT/MDM).

Before I outline the findings concerning which brain areas are most active during conscious processing of visual information, I will briefly describe two methods that are used to study neural activity in visual consciousness, namely visual masking and binocular rivalry. Then I will describe the mechanisms underlying stimulus competition in binocular rivalry as these mechanisms are assumed to be heavily related to the ones at work in visual consciousness.

Concerning methods, the basic principle is similar across research paradigms: Contrasting the neural activity when a stimulus is processed consciously with the neural activity that occurs when a stimulus is processed unconsciously (Baars, 2007). The difficult part is establishing that a comparable stimulus is processed unconsciously. In visual masking this is accomplished through eliminating the visibility of one brief stimulus, referred to as the target, by presenting another brief stimulus, referred to as the mask (Price, 2001). In binocular rivalry on the other hand this is accomplished by presenting different images to each eye. When the images are presented in this manner,

they are presumed to compete for perceptual dominance (Logothetis, et. al., 1996), making only one image is visible at a time (Maier, et. al., 2011).

With regards to the mechanisms underlying stimulus competition in binocular rivalry, it is interesting to note that subjects are unable to wilfully trigger or suppress shifts from one eye to the other as this is a largely involuntary process (Blake & Tong, 2008). However, early visual scientists like Helmholtz (1925) noted that it was possible to delay the shifts with focused attention. In addition to the attentional influence, perceptual dominance is also influenced by factors including familiarity (LoSciuto & Hartley, 1963; Walker, 1978), as well as other affective factors (Alpers, et. al., 2005; Alpers & Pauli, 2006; Alpers & Gerders, 2007), indicating that endogenous attentional factors influence perceptual dominance (Chong & Blake, 2006).

Whether the discrepant monocular patterns compete for awareness due to rivalry between pattern representations or rivalry between monocular pathways has been a subject of debate (Blake, 2001; Tong, 2001). A hybrid-model (Tong, et. al., 2006), integrating views emphasizing low-level reciprocal inhibitory connections with higher level excitatory influences, has to a large degree resolved this disagreement (Blake & Tong, 2008).

This model primarily concerns itself with eye and pattern-based suppression, where eye-based suppression is an endogenous effect and concerns an inhibition of the visual field stemming from one eye, which in turn is an effect of the perceptual systems preference for a single visual stream (Tong, et. al., 2006). By contrast, pattern-based suppression is an exogenous effect and refers to an inhibition of the visual stimuli in

question. In addition to these presumably unconscious processes, the authors also include top down processes to account for attention and perceptual grouping.

Specifically, they hypothesized that the neural basis of eye-based and pattern-based suppression is reciprocal inhibitory connections between monocular neurons (i.e. neurons with a preferred eye of origin) and binocular neurons. By contrast, reciprocal excitatory connections between contralateral areas are considered a viable account for eye based grouping and low-level grouping between monocular neurons with a similar pattern preference (Tong, et. al., 2006). The latter includes grouping that happens interoculary, as well as pattern-based higher level grouping between binocular neurons. Lastly, excitatory feedback projections reduce clutter and are proposed as a mechanism for top down effects of attention and feedback effects of perceptual grouping. This view can account for the mechanisms of competition between exogenous and endogenous factors, with the two of them rivalling their parallel, as well as each other at numerous levels in the visual stream (Blake & Logothetis, 2002; Blake & Tong, 2008). Continuously modulated by feedback effects from perceptual grouping and attentional processes, which in turn influence and are influenced by various states in the observer, both cognitive and affective.

Considered jointly, this theory, with its vast number of reciprocal inhibitory and reciprocal excitatory effects - as well as feedback projections bringing about different kinds of eye-based and pattern-based suppressions and groupings, in addition to attentional effects, fails to present a straightforward picture of where consciousness arises. However, the hybrid theory (Tong, et. al., 2006) on which there is agreement upon among scientists (Blake & Tong, 2008), sits very well with consciousness being an

emergent property of global neuronal networks as argued by the GWT and MDM. It is not as compatible with the notion of consciousness being dependent upon higher-order thoughts, which arise in the frontal lobe. According to the hybrid model, there are various forms of eye-based groupings and pattern-based groupings brought about by inhibitory and excitatory reciprocal connections at both a lower level and a higher level. Hence, a large proportion the causal chain leading up to something becoming conscious appears to have been ascribed to mechanisms located in the lower levels and not the higher-order thoughts, which are located in the prefrontal cortex.

Attributing a fraction of the causal chain leading up to something becoming conscious to lower-level activity in the hybrid model (Tong, et. al., 2006), is also in line with the studies on blindsight that show that discrete lesions to the striate cortex can render a person cortically blind (Ramachandran & Blakeslee, 1998), indicating that the striate cortex is crucial for visual awareness. However, in the striate and prestriate cortex, only 14% of the recorded sites fire at an increased rate when a stimulus is conscious (Gail et. al., 2004; Keliris, et. al., 2010). On the other hand, studies on blind spots have indicated that certain neurons in the striate cortex have receptive fields sensitive to large uniform patterns covering the blind spot, and are more active in conditions eliciting 'filling in' (Komatsu, et. al., 2000). Other studies have indicated that extrastriate neurons are active when conditions eliciting 'filling in' are met (De Weerd, et. al., 1995).

On the other hand, in extrastriate areas, increased activity is only weakly correlated with visual consciousness (Leopold & Logothetis, 1996; Maier, et. al., 2007). Studies on binocular rivalry indicate the combination of lower level inhibitory

connections and higher level excitatory influences which forms the basis of the hybrid model (Tong, et. al., 2006). In line with this, some have theorized that the activity in the prestriate and striate areas reflect perceptual organization mechanisms that give rise to subjective visual consciousness in the temporal lobe where neural activity explicitly represents the conscious content (Logothetis, 1998; Blake & Logothetis, 2002).

Supporting this proposal are studies associating visual consciousness with activity in the inferior temporal cortex and the visual areas of the temporal sulcus (Sheinberg & Logothetis, 1997). Some however argue that the reciprocal connection between these areas and the visual areas of the lateral prefrontal cortex, where neuronal populations are known to respond selectively to faces and other intricate perceptual stimuli, is a basis for inferring that visual consciousness might arise in the lateral prefrontal cortex. This proposal is supported by evidence showing a higher activity during conscious awareness, in the lateral prefrontal cortex, as well as temporal areas (Panagiotarapoulos, et. al., 2012). Lastly, another study has implicated the mid dorsolateral prefrontal cortex (Lau & Passingham, 2006)

As a whole, the findings are consistent with the proposal of visual consciousness being an emergent property of global networks of neuronal populations (Blake & Logothetis, 2002; Panagiotarapoulos, et. al., 2012). Though prefrontal activity is implicated in visual consciousness (Lau & Passingham, 2006), the other evidence reviewed here indicates that temporal areas are correlated with conscious experience as well (Sheinberg & Logothetis, 1997). This suggests that prefrontal areas have a role within a larger neural network, but not that those areas are on their own sufficient for the translation from non-conscious neural processing to phenomenal consciousness.

Discussion of findings

Having reviewed evidence relating to hypotheses A1, A2 and B1, B2, which pertain to questions A (i.e. how is it that at least some organisms (e.g. humans) have subjective experiences?) and B (i.e. what is the connection between subjectivity and the physical systems known as brains?) of the HP, I now discuss the findings further.

Function and Consciousness

Considering the competing hypothesis A1 and A2, that make predictions about the function of consciousness. The evidence reviewed here does not provide grounds for any authorative claims about the function of consciousness. Nor does it present a basis for concluding whether consciousness is best understood in terms of function. However, it does provide a refutation of the denial of such claims. Based on the comparison of the absent qualia/philosophical zombie thought experiments and the research on patients with blindsight, I would argue that whether consciousness can be understood in terms of its function can be informed by studies showing functional and phenomenal dissociation or association. Pointing out the large discrepancy between the imagined philosophical zombie and the real world philosophical zombie indicates that whether there is a functional and phenomenal dissociation or association is an empirical question.

However, if the evidence here did provide authorative grounds for thinking in terms of a functional and phenomenal association, this would not in and of itself falsify

hypothesis A2, which is held by the HOT. This theory is agnostic towards the question of whether consciousness has a function (Lau & Rosenthal), though explicitly presented as non-epiphenomenalist (Rosenthal, 2008). Taking this view into consideration, one could ask whether *any* evidence could falsify this hypothesis. If no evidence could falsify it, one could question whether it should be regarded as a scientific theory at all.

A related theme is conceptual rather than empirical. According to the HOT, conscious awareness is crucially dependent upon a higher-order thought that represents the person in question as being in a particular mental state (Lau & Rosenthal, 2011). Conversely, in the GWT and MDM consciousness is seen as global accessibility (Baars, 2002). This is because global neuronal networks cannot be the cause of consciousness. Rather it is consciousness (Dennett, 2001). HOT however relies on a causal relationship between higher-order thoughts and consciousness. Placing consciousness at the end of a causal chain.

As in disagreement with this stance, Dennett (2001) warns us that it is important to be wary of the anthropocentric perspective that implicitly assumes consciousness is the end product of a causal chain. This is comparable to thinking that apples are the end product of apple trees, Dennett argues, and continues to write that the end product of apple trees are more apple trees. Any other account than a functional one will not avoid this pitfall. This warning echoes Darwin (1838) who wrote, *"Why is thought, being a secretion of the brain more wonderful than gravity a property of matter? It is our arrogance, it is our admiration of ourselves.." (pp. 160-161)* in his notebook.

Neural Activity and Consciousness

The findings bearing on the competing hypothesis B1, and B2 paint a clearer picture than those impacting hypothesis A1 and A2. The evidence supports B2, i.e., consciousness is an emergent property of global neuronal networks (GWT/MDM).

However, there is evidence that activity in prefrontal areas is correlated with the experience of visual consciousness (Lau & Passingham, 2006). But as I have argued earlier, temporal areas are also strongly correlated with the experience of visual consciousness (Sheinberg & Logothetis, 1997). It appears as if the HOT does not sit as well with these findings as the GWT and MDM.

To develop this theme further it is important to draw a distinction between sufficient and necessary conditions for visual consciousness. Some low level visual areas appear to be necessary for visual experience, such as the striate cortex where a large bilateral lesion can make a person cortically blind (Ramachandran & Blakeslee, 1998). These areas however, might not in themselves be sufficient. If we return to the HOT, it claims that a state is conscious due to representational relation between a higher-order thought and a mental state (Rosenthal, 1990). This higher-order thought crucially arises in the prefrontal cortex (Lau & Rosenthal, 2011). According to this view an intact prefrontal cortex is necessary for visual consciousness.

In line with this, Lau and Rosenthal (2011) argue that many studies link frontal lobe lesions to disturbances in various responses to visual stimuli (e.g. Barcelo, et. al., 2000; Guitton & Buchtel, 1985; Kennard, 1939). However , these lesions do not abolish awareness completely. The only study referred to by Lau and Rosenthal (2011) showing

this effect is a study on monkeys, where the parietal lobe has been removed in addition to the frontal lobe (e.g. Nakamura & Mishkin, 1986). Another study that these authors refer to, shows impaired meta-cognitive ability (e.g. poorer stimulus discrimination abilities and lower reported levels of visual awareness) in visual awareness tasks. This effect was seen with lesions to the dorsolateral prefrontal cortex (Rounis et. al., 2010) that were temporarily induced by transcranial magnetic stimulation (TMS). This is interesting, although not an indicative for prefrontal areas being a necessary condition for visual awareness. The findings presented by Rounis et. al., (2010) could just as well be interpreted as support for hypothesis B2, something that would be compatible with the evidence pointing to the relevance of other brain areas as well for visual consciousness.

Summarising this discussion, with the current evidence it is presently impossible to decisively conclude on the nature of the connection between subjectivity and brains. However, hypothesis B2 show more promise than hypothesis B1 with this regards to this.

Implications of Findings

Any paper that discusses theoretical solutions to the HP should also include some limitations. Firstly, the evidence reviewed here does not close the explanatory gap that is addressed by the HP. This was not the intention of this paper, which aims instead to provide an evaluation of some influential theoretical approaches to the HP.

Consistent with this, there are limitations to this paper that were evident from the outset as it rests on the assumption that a naturalistic account of consciousness might be possible. Contrary to this position, there are philosophers and scientists who think that a naturalistic account is impossible. These authors have metaphysical beliefs about consciousness ranging from materialism (e.g. Pinker, 1997) to mysterianism (e.g. McGinn, 1991). When the foundation of consciousness is a subject to debate, the accounts of of it could be wrong at a more fundamental level than what was within the scope of this paper.

However - according to Van Gulick (2007), various forms of functionalism are the most commonly held positions among philosophers with regards to the nature of mental states. Multiple sources also make reference to a surge of interest in the scientific study of consciousness (e.g. Blackmore, 2010; Dennett, 2001; Mind-Science Foundation, 2013; Velmans & Schneider, 2007). Considered jointly, this makes whether consciousness can be explained scientifically an open question.

Though this paper proceeded from that assumption, the theories discussed in this paper remain a small subset of scientific theories of consciousness. A fact that limits the application of the points that are argued here.

In line with this, I have not proved that the solutions provided by the GWT or MDM are correct. However, this paper outlines some central problems with the HOT, indicating that the theoretical solution outlined by this theory is less compatible with empirical data than the theoretical solution provided by the GWT and the MDM. Based on this, I have argued that there are good reasons not to adopt a priori assumptions of

thinking in terms of phenomenal and functional dissociation. I have also argued that there are good reasons to think that consciousness is linked to activity in global neuronal networks, rather than linked to one specific area.

A further clarification of the position I have argued is Dennett's (2001) warning about the anthropocentric fallacy associated with viewing consciousness as the end of a causal chain. Activity in global neuronal networks cannot be the cause of consciousness. Rather it *is consciousness*. HOT which does not explain consciousness in terms of function appears to be vulnerable to this fallacy.

However, a number of philosophers claim that functional explanations are unable to address phenomenal consciousness or qualia (e.g. Chalmers, 2004; Block, 1980). Some of the arguments against functionalism are thought experiments which make strong a priori claims (Van Gulick, 2007). In this paper I have argued against such claims and suggested that a functional and phenomenal dissociation is an empirical question.

This argument was intended as a critique of certain empirical claims rather than as a direct contribution to the qualia-functionalism debate, which is philosophical rather than scientific and therefore beyond the scope of this paper. However – some aspects of this debate should be considered in relation to the theories presented in this paper. The interpretive dimensions on which this dispute plays out varies between: (1) how functionalism should be interpreted, (2) which real features that can be ascribed to qualia, (3) and what the appropriate standard for explanatory success is (Van Gulick, 2007). All of the theories discussed in this paper have their own assumptions built into

them in terms of the interprative dimensions above - something that was touched upon when the theories' stance on the HP was discussed. In line with this, the theories might meet their own standard for explanatory success without meeting standards that are deemed appropriate by other theories.

Hence, the theories discussed here might have assumptions and standards built into them that are simply incompatible with each other. However, all the theories discussed here purport to be scientific theories that make specific empirical predictions. If this is the case, the problem of incompatability should not arise. Should the differences between the theories discussed here be the result of an incompatability, this would not in itself make the discussion and its conclusions erroneous. On the contrary, the testing of different assumptions and standards could contribute to clarify and elaborate on which of these assumptions that will enable scientists to capture the elusive concept of phenomenal consciousness while avoiding violation of scientific criteria.

With regards to this, I have argued that understanding consciousness as an emergent property of global access/global neuronal networks shows promise. As opposed to understanding consciousness as crucially dependent upon a representational relation between a higher-order thought and a mental state.

On the other hand, one might argue that these accounts of consciousness are far removed from the 'raw feels' of viewing for example a birch forest, alive with sunlight and bright yellow colours during fall. The reply from Dennett (2001), who will serve as an example of a functionalist (note that there are many other versions), would be that the subjective qualities of this experience only accessible to the individual must be

broken down in complex dispositional traits distributed in space and time in the brain and analyzed in a functional manner.

Regardless of how matter-of-fact that explanation sounds in comparison with what it is trying to explain, it is an explanation, or at least a recipe for an explanation. Though the explanatory gap still exists, this perspective on consciousness provides an opening for examining and understanding parts of the phenomenon.

While the theories discussed here provided an opening for understanding the phenomenon, there are many theoretical solutions that were not included in this paper. These theories might fit the evidence equally as well, or maybe even better than the theories discussed here. Consistent with this, it is reasonable to assume that theories of consciousness will become more detailed, when better brain imaging techniques and research paradigms are discovered.

Considering the relatively low resolution of brain imaging techniques, a suggestive comparison is the group of astronomers led by Lowell (1906). At the turn of the 19th century they were convinced that there were canals on Mars, indicating the existence or former existence of intelligent life on this planet (Lowell, 1904). They spent considerable time gazing through telescopes, and drew elaborate maps of the canal network (Sagan & Fox, 1975). For a time, canals on mars were the established opinion, until perceptual scientists pointed to optical illusion effects, and the fact that it is generally easier to see what you want to see with coarsely grained data (MariaLane, 2006). This issue was resolved in favor of no canals on mars in 1965, with the images from Mariner 4 (Gerstbach, 2003).

Studying the neural correlates of consciousness is of course an important project, but so was establishing whether or not there were canals on mars. The point is not to ridicule Lowell's research group, as their project was important at that time. Rather, my goal is to illustrate that coarsely grained data warrants a cautious interpretation; strong claims about consciousness are at this point likely to be false.

References

Alpers, G. W., Ruhleder, M., Walz, N., Mühlberger, A., & Pauli, P. (2005). Binocular rivalry between emotional and neutral stimuli: A validation using fear conditioning and EEG. *International Journal of Psychophysiology*, *57*, 25-32. doi:http://dx.doi.org/10.1016/j.ijpsycho.2005.01.008

Alpers, G., & Pauli, P. (2006). Emotional pictures predominate in binocular rivalry. *Cognition & Emotion,* 20, 596-607. doi:10.1080/02699930500282249

Alpers, G. W., & Gerdes, A. (2007). Here is looking at you: emotional faces predominate in binocular rivalry. *Emotion*, 7, 495. doi: 10.1037/1528-3542.7.3.495

Armstrong, D. (1968) *A Materialist Theory of the Mind,* London: Routledge

Baars, B.J. (1988) *A Cognitive Theory of Consciousness.* New York: Cambridge University Press.

Baars, B.J:, (2002) The conscious access hypothesis: origins and recent evidence,

*Trends in Cognitive Sciences,*1, 47-52 doi:http://dx.doi.org/10.1016/S1364-6613(00)01819-2

Baars, B.J. (2007) *Global Workspace Theory of consciousness,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 236-246, Malden : Blackwell publishing

Barcelo, F., Suwazono, S., & Knight, R. T. (2000). Prefrontal modulation of visual processing in humans. *Nature neuroscience*, *3*, 399-403, doi:10.1038/73975

Blackmore, S. (2010) *Consciousness, an Introduction,* London: Hodder education

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and mind*, *2*, 5-38. doi: 10.1023/A:1017925416289

Blake, R., & Logothetis, N. K., (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13-21. doi:10.1038/nrn701

Blake R. & Tong F., (2008) Binocular rivalry. *Scholarpedia, the peer reviewed, open access encyclopedia* 3(12):1578.

Block, N. (1980). Troubles with functionalism. Block, N., (Ed) 1980, In *Readings in the Philosophy of Psychology*, *1*, 268-305., Cambridge, MA: Harvard University Press

Block, N. (1995) On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227–47. doi:http://dx.doi.org/10.1017/S0140525X00038346

Byrne, A., (1997) Some like it HOT: Consciousness and Higher-Order Thoughts. *Philosophical Studies,* 86, pp. 103-129. doi:10.1023/A:1017959107565

Carruthers, P., (2003) *Phenomenal Consciousness: A Naturalistic Theory.* Cambridge: Cambridge University Press

Carruthers, P. (2007) *Higher-Order Theories of Consciousness,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell Companion to Consciousness,* 277-286, Malden: Blackwell publishing

Chong, S. C., & Blake, R. (2006). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision research*, *46*, 1794-1803. doi: http://dx.doi.org/10.1016/j.visres.2005.10.031

Chalmers, D.J., (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2(3)***,** 200–219

Chalmers, D.J., (2004). Imagination, indexicality, and intensions. *Philosophy and Phenomenological Research*, *68*(1), 182-190.

Chalmers, D.J., (2007) *Naturalistic Dualism,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 359-368 Malden: Blackwell publishing

Churchland, P. M., & Churchland, P. S. (1998). *Recent work on consciousness,*

philosophical, theoretical, empirical, In, Churchland, P. M., & Churchland, P. S. (1998). *On the contrary: Critical essays, 1987-1997*. Cambridge MA: MIT Press.

Cowey, A., & Stoerig, P. (1991). The neurobiology of blindsight. *Trends in neurosciences*, *14*, 140-145. doi: http://dx.doi.org/10.1016/0166-2236(91)90085-9

Culham, J. C., Danckert, S. L., De Souza, J. F., Gati, J. S., Menon, R. S., & Goodale,. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research*, *153*, 180-189. doi: 10.1007/s00221-003-1591-5

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences* (2) 263-275.

Darwin, C.R., (1838) *Notebook C*, Archival source, author reasonably certain, date not stated on document. Age is stated. Date is inferred from year of birth (1809). As retrieved on 24.11.13 http://www.darwinproject.ac.uk/wp-content/uploads/2012/08/Excerpts-from-Notebook-C-copy.pdf

Dean, P., Redgrave, P., & Westby, G. W. M. (1989). Event or emergency? Two response systems in the mammalian superior colliculus. *Trends in neurosciences*, *12*, 137-147. doi: http://dx.doi.org/10.1016/0166-2236(89)90052-0

Dehaene, S., & Changeux, J. P. (2000). Reward-dependent learning in neuronal networks for planning and decision making. *Progress in brain research*, *126*, 217-229. doi:http://dx.doi.org/10.1016/S0079-6123(00)26016-0

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature neuroscience*, *4*, 752-758. doi:10.1038/89551

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*,1-37. doi:http://dx.doi.org/10.1016/S0010-0277(00)00123-2

Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences*, *100*, 8520-8525. doi:10.1073/pnas.1332574100

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences*, *10*, 204-211. doi:http://dx.doi.org/10.1016/j.tics.2006.03.007

Dennett, D.C., (1988) *Quining Qualia,* In, Marcel, A.J., & Bisiach E., (Eds.) *Consciousness in modern science,* pp 42-77, Clarendon Press: Oxford,.

Dennett, D.C., (1989) *The Intentional Stance,* Cambridge MA: MIT Press

Dennett, D.C., (1991) *Consciousness Explained,* London: Penguin Books

Dennett, D.C., (1995). The unimagined preposterousness of zombies. *Journal of Consciousness studies*, *2(4)*, 322-326.

Dennett, D.C., (1996a). *Consciousness: More like fame than television*. Delivered at a

Munich conference, German paper published as, *Bewusstsein hat mehr mit Ruhm als mit Fernsehen zu tun,* In Maar, C., Pöppel, E., & Christaller T.,(Eds.), *Die Technik auf dem Weg zur Seele*. Munich: Rowhoult

Dennett, D. C. (1996b). Facing backwards on the problem of consciousness. *Journal of Consciousness Studies*, *3*(1), 4-6.

Dennett, D.C., (2001). Are we explaining consciousness yet?. *Cognition*, *79*, 221-237 doi: http://dx.doi.org/10.1016/S0010-0277(00)00130-X

Dennett, D.C, & Akins, K., (2008) Multiple Drafts Model. *Scholarpedia, the peer reviewed, open access encyclopedia* 3(4):4321.

Descartes, R. (1996). *Descartes: Meditations on first philosophy, 1641: With selections from the objections and replies*. Cambridge: University Press.

Droege, P. (2005) Higher-order theories of Consciousness, *Internet Encyclopedia of Philosophy, a peer reviewed academic resource*, http://www.iep.utm.edu/consc-hi/

Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and cognition*, *8*, 128-151. doi: http://dx.doi.org/10.1006/ccog.1999.0389

Fendrich, R., Wessinger, C. M., & Gazzaniga, M. S. (1992). Residual vision in a scotoma: implications for blindsight. *Science*, *258*, 1489-1491. doi: 10.1126/science.1439839

Gail, A., Brinksmeyer, H. J., & Eckhorn, R. (2004). Perception-related modulations of local field potential power and coherence in primary visual cortex of awake monkey during binocular rivalry. *Cerebral Cortex*, *14*, 300-313. doi:10.1093/cercor/bhg129

Gerstbach, G. (2003). Mars channel observations 1877-90, compared with modern Orbiter data. *Publications of the Astronomical Observatory of Belgrade*, *75*, 347-354.

Gray, J. (2004). *Consciousness: Creeping up on the hard problem*. Oxford: Oxford University Press.

Grush, R., & Churchland P.S., (1995) Gaps in Penrose's toilings. *Journal of Consciousness Studies,(*2) 1, 10-29

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research*, *58*, 455-472. doi:http://dx.doi.org/10.1007/BF00235863

Hameroff, S.R., & Penrose, R. (1996) Orchestrated reduction of quantum coherence in the brain micro-tubules: A model for consciousness. *Mathematics and Computers in Simulation,* 40, 453-480 doi:http://dx.doi.org/10.1016/0378-4754(96)80476-9

Hameroff, S. (1998) Anesthesia, consciousness and hydrophobic pockets, a unitary

quantum hypothesis of anesthetic action, *Toxicology Letters*, 100-101, 31-39. doi: http://dx.doi.org/10.1016/S0378-4274(98)00162-3

Harnad, S. (1994). Why and how we are not zombies. *Journal of Consciousness Studies*, *1*(2), 164-167.

Helmholtz, von., H., (1925) *Treatise on Physiological Optics, vol, 3, the optical society of America's Southall translation,* Rochester, NY: Optical society of America

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, *10*, 438-443. doi: 10.1111/1467-9280.00183

Hodges, J. R., Spatt, J., & Patterson, K. (1999). "What" and "how": Evidence for the dissociation of object knowledge and mechanical problem-solving skills in the human brain. *Proceedings of the National Academy of Sciences*, *96*, 9444-9448. doi:10.1073/pnas.96.16.9444

Howell R.J., & Alter T., (2009) Hard problem of consciousness. *Scholarpedia, the peer reviewed, open access encyclopedia* 4(6):4948.

Keliris, G. A., Logothetis, N. K., & Tolias, A. S. (2010). The role of the primary visual cortex in perceptual suppression of salient visual stimuli. *The Journal of Neuroscience*, *30*, 12353-12365. doi:10.1523/ JNEUROSCI.0677-10.2010

Kennard, M. A. (1939). Alterations in response to visual stimuli following lesions of frontal lobe in monkeys. *Archives of Neurology and Psychiatry*, *41*, 1153. doi: 10.1001/archneurpsyc. 1939.02270180081008.

Komatsu, H., Kinoshita, M., & Murakami, I. (2000). Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *The Journal of Neuroscience*, *20*, 9310-9319.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, *103*(49), 18763-18768. doi: 10.1073/pnas.0607716103

Lau, H., & Rosenthal, D.M., (2011) Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences,* 15, 365-373 doi:http://dx.doi.org/10.1016/j.tics.2011.05.009

Leopold, D. A., & Logothetis, N. K. (1996). Activity Changes in Early Visual-Cortex Reflect Monkeys Percepts During Binocular-Rivalry. *Nature*, *379*, 549-553. doi: 10.1038/379549a0

Libet, B. (1996). Solutions to the hard problem of consciousness. *Journal of Consciousness Studies*, *3*, 33-35.

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What Is Rivaling During Binocular-Rivalry. *Nature*, *380*, 621-624. doi:10.1038/380621a0

Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*, 1801-1818. doi: 10.1098/rstb.1998.0333

LoSciuto, L. A., & Hartley, E. L. (1963). Religious affiliation and open-mindedness in binocular resolution. *Perceptual and motor skills*. 17, 427-430,

Lowell, P., (1904). The cartouches of the canals of Mars. *Lowell Observatory Bulletin*,

1, 59-86.

Lowell, P., (1906). *Mars and its Canals*. London: Macmillan & co.,

Lycan, W., (1995) *Consciousness.* Cambridge, MA: MIT Press

Lyon, D. C., Nassi, J. J., & Callaway, E. M. (2010). A disynaptic relay from superior colliculus to dorsal stream visual cortex in macaque monkey. *Neuron*, *65*, 270-279, doi: http://dx.doi.org/10.1016/j.neuron.2010.01.003

Maier, A., Logothetis, N. K., & Leopold, D. A. (2007). Context-dependent perceptual modulation of single neurons in primate visual cortex. *Proceedings of the National Academy of Sciences*, *104*, 5620-5625.doi:10.1073/pnas.0608489104

Maier, A., Panagiotaropoulos, T. I., Tsuchiya, N., & Keliris, G. A. (2011). Introduction to Research Topic Binocular Rivalry: A Gateway to Studying Consciousness. *Frontiers in Human Neuroscienc*e, 6, 263-263. doi:10.3389/fnhum.2012.00263

MariaLane, K. D. (2006). Mapping the Mars canal mania: Cartographic projection and the creation of a popular icon. *Imago Mundi*, *58*, 198-211. doi: 10.1080/03085690600687255

McGinn, C. (1991). *The problem of consciousness: Essays towards a resolution.* Oxford: Blackwell.

McGinn, C. (2000). *The mysterious flame: Conscious minds in a material world.* New York: Basic Books

McGinn, C. (2004) *Consciousness and Its Objects* Oxford: Oxford University press.

Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. In Hameroff, S.,

Kaszniak, A.W., Chalmers, D.J., (Eds.) *Toward a Science of Consciousness III, the third Tucson discussion and debates.* Cambridge: MIT Press

Milner, D. (1998). Insights into blindsight. *Trends in cognitive sciences*, *2*, 237-238. doi: http://dx.doi.org/10.1016/S1364-6613(98)01185-1

Mind-Science Foundation (2013) *The Scientific Study of Consciousness,* as retrieved 03.11.13. http://www.mindscience.org/research/consciousness-intro

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in neurosciences*, *6*, 414-417. doi: http://dx.doi.org/10.1016/0166-2236(83)90190-X

Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences*, *96*, 1680-1685. doi: 10.1073/pnas.96.4.1680

Mormann F., & Koch C., (2007) Neural correlates of consciousness. *Scholarpedia, the peer reviewed, open access encyclopedia, 2(12):1740.*

Nagel, T., (1974) What is it like to be a bat?, *Philosophical Review,* 83, 435-450

Nakamura, R. K., & Mishkin, M. (1986). Chronic 'blindness' following lesions of nonvisual cortex in the monkey. *Experimental brain research*, *63*, 173-184. doi: http://dx.doi.org/10.1007/BF00235661

Panagiotaropoulos, T. I., Deco, G., Kapoor, V., & Logothetis, N. K. (2012). Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron*, *74*, 924-935.doi:http://dx.doi.org/10.1016/j.neuron.2012.04.013

Penrose, R., (1994) *Shadows of the Mind,* New York: Oxford University Press

Persaud, N., & Lau, H. (2008). Direct assessment of qualia in a blindsight participant. *Consciousness and cognition*, *17*, 1046-1049. doi:http://dx.doi.org/10.1016/j.concog.2007.10.001

Pinker, S. (1997) *How the Mind Works* New York: Norton

Putnam, H., (1994) Review of Roger Penrose, *Shadows of the Mind, New York Times Book Review,* November 20. http://www.nytimes.com/books/97/04/27/nnp/17540.html, as retrieved 19.11.2013

Price, M.C., (2001). Now you see it, now you don't. Preventing consciousness with visual masking. In P.G. Grossenbacher (ed.), *Finding Consciousness in the Brain: A Neurocognitive Approach (Advances in Consciousness Research, 8),* pp. 25-60. Amsterdam: John Benjamins.

Prinz, J. (2007) *The Intermediate level theory of Consciousness,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 347-60, Malden : Blackwell publishing

Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*, 699-702. doi: 10.1038/350699a0

Ramachandran, V. S., Gregory, R. L., & Aiken, W. (1993). Perceptual fading of visual texture borders. *Vision Research*, *33*, 717-721. doi: http://dx.doi.org/10.1016/0042-6989(93)90191-X

Ramachandran, V. S., & Blakeslee, S. (1998). Phantoms in the brain: Probing the mysteries of the human mind. *London, UK: Fourth Estate*.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, *3*, 261-270. doi:10.1038/nrn783

Rosenthal, D. M. (1990). *A theory of consciousness* (pp-729-753). Issued as Report 40/1990, Center for Interdisciplinary Research (ZiF), Universitat Bielefeld, Germany

Rosenthal, D.M., (2000) Consciousness, Content, and Metacognitive Judgments, *Consciousness and Cognition*, 9, 203-214

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia*, *46*, 829-840.doi:http://dx.doi.org/10.1016/j.neuropsychologia.2007.11

Rosenthal, D.M, & Weisberg J. (2008) Higher-order theories of consciousness. *Scholarpedia, the peer reviewed, open access encyclopedia* 3(5):4407

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, *1*, 165-175. doi: 10.1080/17588921003632529

Rowlands, M. (2007) *Mysterianism,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 335-346, Malden : Blackwell publishing

Sagan, C., & Fox, P. (1975). The canals of Mars: an assessment after Mariner 9. *Icarus*,

25, 602-612.doi: http://dx.doi.org/10.1016/0019-1035(75)90042-1

Seager, W., & Bourget, D., (2007) *Representationalism about consciousness*, In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 261-277, Malden: Blackwell publishing

Searle, J., (2007) *Biological Naturalism,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 325-334, Malden: Blackwell publishing

Sahibzada, N., Dean, P., & Redgrave, P. (1986). Movements resembling orientation or avoidance elicited by electrical stimulation of the superior colliculus in rats. *The Journal of neuroscience*,3(*6*), 723-733.

Shear, J. (Ed.). (1999). *Explaining consciousness: The hard problem*. Cambridge MA: MIT Press.

Shmuelof, L., & Zohary, E. (2005). Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron*, *47*, 457-470. doi: http://dx.doi.org/10.1016/j.neuron.2005.06.034

Stapp, H. (2007) *Quantum Mechanical Theories of Consciousness*, In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 300-312, Malden: Blackwell publishing

Schneider, S. (2007) *Daniel Dennett on the Nature of Consciousness,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 313-324, Malden: Blackwell publishing

Shanahan, M., & Baars, B.J. (2005). Applying Global Workspace Theory to the frame problem. *Cognition*, *98*, 157-176. doi:10.1016/jcognition.2004.11.007

Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences*, *94*, 3408-3413. doi: 10.1073/pnas.94.7.3408

Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind*, *2*, 55-83. doi:10.1023/A:1017942718744

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in cognitive sciences*, *10*, 502-511. doi: http://dx.doi.org/10.1016/j.tics.2006.09.003

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness. *Annals of the New York Academy of Sciences*, *1124*, 239-261. doi:10.1196/annals.1440.004

Tononi, G. (2007) *The Information Integration Theory of Consciousness,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 287-299, Malden: Blackwell publishing

Tye, M., (2013) *Qualia*, In Zalta E.N., (ed.) (2013) *The Stanford Encyclopedia of Philosophy* (Summer edition), As retrieved, 24.11.13 on http://plato.stanford.edu/archives/fall2013/entries/qualia/

University of Massachusetts Amherst Libraries (2013) *Top 50 psychology journals, ranked by JCR impact factor,* as retrieved, 03.11.13, http://guides.library.umass.edu/content.php?pid=52227&sid=383170

Valyear, K. F., Culham, J. C., Sharif, N., Westwood, D., & Goodale, M. A. (2006). A

double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia*, 44, 218-228. doi: http://dx.doi.org/10.1016/j.neuropsychologia.2005.05

Van Gulick, R. (2007) *Functionalism and Qualia* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* 381-395 Malden: Blackwell publishing

Van Gulick, R. (2011) *Consciousness*, In Zalta, E.N., (Ed.) (2013) *The Stanford Encyclopedia of Philosophy* (Summer Edition 2013), as retrieved 24.11.13 on http://plato.stanford.edu/archives/sum2011/entries/consciousness/

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of consciousness studies*, *3*(4), 330-349.

Velmans, M. (2007), *Dualism, reductionism and reflexive monism,* In Velmans M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,*346-358 Malden: Blackwell publishing

Velmans, M., & Schneider S., (Eds.) (2007) *Blackwell companion to Consciousness,* Malden: Blackwell publishing

Walker, P. (1978). Binocular rivalry: Central or peripheral selective processes?. *Psychological Bulletin*, *85*, 376-389. doi: 10.1037/0033-2909.85.2.376

Weerd, P. D., Gattass, R., Desimone, R., & Ungerleider, L. G. (1995). Responses of cells in monkey visual cortex during perceptual filling-in of an artificial scotoma. *Nature*, *377*, 731-734. doi:10.1038/377731a0

Weiskrantz, L., & Weiskrantz, L. (1986). *Blindsight: A case study and implications*

(Vol. 12). Oxford: Clarendon Press.

Zeki, S., & Bartels, A. (1999). Toward a theory of visual consciousness. *Consciousness and cognition*, *8*, 225-259. doi: http://dx.doi.org/10.1006/ccog.1999.0390